

# Stat 602 Homework 2

*Kiegan Rice and Nate Garton*

*Due 3/4/2019*

## Problem 1

### Problem 4.3

### Problem 4.4

#### Part a

```
## k-Nearest Neighbors
##
## 146 samples
## 9 predictor
## 2 classes: '1', '2'
##
## Pre-processing: centered (9), scaled (9)
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 131, 132, 132, 132, 131, 131, ...
## Resampling results across tuning parameters:
##
## k Accuracy Kappa
## 1 0.8147619 0.6287535
## 2 0.7800000 0.5600949
## 3 0.7876190 0.5758517
## 4 0.7871429 0.5753205
## 5 0.7809524 0.5613518
## 6 0.7671429 0.5365818
## 7 0.7742857 0.5494650
## 8 0.7666667 0.5358735
## 9 0.7938095 0.5895225
## 10 0.8009524 0.6045257
## 11 0.7533333 0.5102489
## 12 0.7190476 0.4432222
## 13 0.7809524 0.5649154
## 14 0.7609524 0.5259698
## 15 0.7614286 0.5270191
## 16 0.7614286 0.5249078
## 17 0.7142857 0.4355299
## 18 0.7147619 0.4364318
## 19 0.7414286 0.4881319
## 20 0.7419048 0.4883815
##
## Accuracy was used to select the optimal model using the largest value.
## The final value used for the model was k = 1.
```

Based on 10-fold cross-validation, the accuracy is estimated to be highest when  $k = 4$ .

#### Part b

The classification rule  $\hat{y} = 2I[t(x) \geq k/2] + I[t(x) < k/2]$  is equivalent to  $\hat{y} = 2I[t(x)/k \geq 1/2] + I[t(x)/k < 1/2]$ . Here,  $t(x)/k$  is an estimate of  $E[\text{type} = 2|x] = P(\text{type} = 2|x)$ . Thus, this is directly an estimate of a posterior class probability which accounts for the prior class probabilities. This means that no modification of kNN needs to be made to account for differing prior probabilities.

### Problem 5.1

The following are decompositions using the matrix

$$X = \begin{bmatrix} 2 & 4 & 7 & 2 \\ 4 & 3 & 5 & 5 \\ 3 & 4 & 6 & 1 \\ 5 & 2 & 4 & 2 \\ 1 & 3 & 4 & 4 \end{bmatrix}$$

#### Part a

The QR decomposition of  $X$  is:

$$Q = \begin{bmatrix} -0.27 & 0.57 & 0.77 & 0.04 \\ -0.54 & -0.07 & -0.13 & 0.56 \\ -0.4 & 0.37 & -0.35 & -0.71 \\ -0.67 & -0.5 & 0.1 & -0.13 \\ -0.13 & 0.53 & -0.5 & 0.4 \end{bmatrix}, \quad R = \begin{bmatrix} -7.42 & -6.07 & -10.25 & -5.53 \\ 0 & 4.15 & 5.99 & 2.28 \\ 0 & 0 & 1.06 & -1.23 \\ 0 & 0 & 0 & 3.57 \end{bmatrix}$$

Thus, the basis for  $C(X)$  will be the columns of the matrix  $Q$ .

The Singular Value Decomposition of  $X$  is:

$$U = \begin{bmatrix} -0.5 & 0.53 & 0.16 & -0.66 \\ -0.5 & -0.59 & 0.13 & -0.11 \\ -0.46 & 0.49 & -0.25 & 0.65 \\ -0.39 & -0.33 & -0.68 & -0.09 \\ -0.37 & -0.17 & 0.65 & 0.35 \end{bmatrix}, \quad D = \begin{bmatrix} 16.58 & 0 & 0 & 0 \\ 0 & 3.78 & 0 & 0 \\ 0 & 0 & 3.38 & 0 \\ 0 & 0 & 0 & 0.55 \end{bmatrix},$$

$$V = \begin{bmatrix} -0.4 & -0.43 & -0.8 & 0.12 \\ -0.44 & 0.3 & 0.18 & 0.83 \\ -0.71 & 0.45 & 0.04 & -0.54 \\ -0.38 & -0.72 & 0.57 & -0.06 \end{bmatrix}$$

Thus, the basis for  $C(X)$  will be the columns of the matrix  $U$ .

#### Part b

We can find the eigen (spectral) decomposition of  $X'X$  by calculating the eigenvalues,  $D^2$ , and the eigenvectors, which will be the columns of  $V$ :

$$D^2 = \begin{bmatrix} 274.94 \\ 14.32 \\ 11.44 \\ 0.3 \end{bmatrix}, \quad V = \begin{bmatrix} -0.4 & -0.43 & -0.8 & 0.12 \\ -0.44 & 0.3 & 0.18 & 0.83 \\ -0.71 & 0.45 & 0.04 & -0.54 \\ -0.38 & -0.72 & 0.57 & -0.06 \end{bmatrix}$$

We can find the eigen (spectral) decomposition of  $XX'$  by calculating the eigenvalues,  $D^2$ , and the eigenvectors, which will be the columns of  $U$  (rather than  $V$  as for  $X'X$ ):

$$D^2 = \begin{bmatrix} 274.94 \\ 14.32 \\ 11.44 \\ 0.3 \end{bmatrix}, \quad V = \begin{bmatrix} -0.5 & 0.53 & 0.16 & -0.66 \\ -0.5 & -0.59 & 0.13 & -0.11 \\ -0.46 & 0.49 & -0.25 & 0.65 \\ -0.39 & -0.33 & -0.68 & -0.09 \\ -0.37 & -0.17 & 0.65 & 0.35 \end{bmatrix}$$

### Part c

The best  $rank = 1$  approximation to  $X$  will be  $X^{*1} = U_1 \text{diag}(d_1) V_1'$ , where  $U_1$  is a matrix with the first column of  $U$ , and  $V_1$  is a matrix with the first column of  $V$ :

$$\begin{aligned} X^{*1} = U_1 \text{diag}(d_1) V_1' &= \begin{bmatrix} -0.5 \\ -0.5 \\ -0.46 \\ -0.39 \\ -0.37 \end{bmatrix} \times [16.58] \times \begin{bmatrix} -0.4 & -0.44 & -0.71 & -0.38 \end{bmatrix} \\ &= \begin{bmatrix} 3.35 & 3.61 & 5.89 & 3.11 \\ 3.38 & 3.64 & 5.94 & 3.14 \\ 3.08 & 3.31 & 5.4 & 2.85 \\ 2.63 & 2.83 & 4.62 & 2.43 \\ 2.45 & 2.64 & 4.31 & 2.27 \end{bmatrix} \end{aligned}$$

The best  $rank = 2$  approximation to  $X$  will be  $X^{*2} = U_2 \text{diag}(d_1, d_2) V_2'$ , where  $U_2$  is a matrix with the first two columns of  $U$ , and  $V_2$  is a matrix with the first two columns of  $V$ :

$$\begin{aligned} X^{*2} = U_2 \text{diag}(d_1, d_2) V_2' &= \begin{bmatrix} -0.5 & 0.53 \\ -0.5 & -0.59 \\ -0.46 & 0.49 \\ -0.39 & -0.33 \\ -0.37 & -0.17 \end{bmatrix} \times \begin{bmatrix} 16.58 & 0 \\ 0 & 3.78 \end{bmatrix} \times \begin{bmatrix} -0.4 & -0.44 & -0.71 & -0.38 \\ -0.43 & 0.3 & 0.45 & -0.72 \end{bmatrix} \\ &= \begin{bmatrix} 2.49 & 4.2 & 6.78 & 1.66 \\ 4.35 & 2.97 & 4.95 & 4.75 \\ 2.28 & 3.86 & 6.23 & 1.51 \\ 3.16 & 2.46 & 4.07 & 3.33 \\ 2.73 & 2.44 & 4.02 & 2.75 \end{bmatrix} \end{aligned}$$

### Part d

Note that

$$\tilde{X} = \begin{bmatrix} -1 & 0.8 & 1.8 & -0.8 \\ 1 & -0.2 & -0.2 & 2.2 \\ 0 & 0.8 & 0.8 & -1.8 \\ 2 & -1.2 & -1.2 & -0.8 \\ -2 & -0.2 & -1.2 & 1.2 \end{bmatrix}$$

The Singular Value Decomposition of  $\tilde{X}$  is:

$$\tilde{U} = \begin{bmatrix} -0.57 & 0.17 & 0.29 & -0.6 \\ 0.51 & 0.11 & 0.7 & 0.21 \\ -0.5 & -0.25 & -0.09 & 0.69 \\ 0.34 & -0.69 & -0.32 & -0.33 \\ 0.22 & 0.65 & -0.57 & 0.03 \end{bmatrix}, \quad \tilde{D} = \begin{bmatrix} 3.83 & 0 & 0 & 0 \\ 0 & 3.38 & 0 & 0 \\ 0 & 0 & 2.01 & 0 \\ 0 & 0 & 0 & 0.48 \end{bmatrix}$$

$$\tilde{V} = \begin{bmatrix} 0.34 & -0.81 & 0.45 & 0.18 \\ -0.37 & 0.18 & 0.26 & 0.88 \\ -0.58 & 0.03 & 0.68 & -0.45 \\ 0.64 & 0.56 & 0.52 & 0 \end{bmatrix}$$

Therefore, the principal component directions will be the columns of  $\tilde{V}$ , the principal components will be the inner products  $z_j = \langle x_i, v_j \rangle$ , the columns of:

$$z = \begin{bmatrix} -2.19 & 0.56 & 0.57 & -0.29 \\ 1.95 & 0.39 & 1.4 & 0.1 \\ -1.91 & -0.84 & -0.18 & 0.33 \\ 1.3 & -2.32 & -0.65 & -0.16 \\ 0.85 & 2.21 & -1.14 & 0.02 \end{bmatrix}$$

The “loadings” of the first principal component are the first column of the matrix  $\tilde{V}$ ,

$$\tilde{V}_1 = \begin{bmatrix} 0.34 \\ -0.37 \\ -0.58 \\ 0.64 \end{bmatrix}$$

### Part e

The best  $rank = 1$  approximation to  $\tilde{X}$  will be  $\tilde{X}^{*1} = \tilde{U}_1 \text{diag}(\tilde{d}_1) \tilde{V}_1'$ , where  $\tilde{U}_1$  is a matrix with first column of  $\tilde{U}$  and  $\tilde{V}_1$  is a matrix with the first column of  $\tilde{V}$ :

$$\tilde{X}^{*1} = \tilde{U}_1 \text{diag}(\tilde{d}_1) \tilde{V}_1' = \begin{bmatrix} -0.57 \\ 0.51 \\ -0.5 \\ 0.34 \\ 0.22 \end{bmatrix} \times [3.83] \times [0.34 \quad -0.37 \quad -0.58 \quad 0.64]$$

$$= \begin{bmatrix} -0.75 & 0.81 & 1.26 & -1.41 \\ 0.67 & -0.72 & -1.12 & 1.26 \\ -0.66 & 0.71 & 1.1 & -1.23 \\ 0.45 & -0.48 & -0.75 & 0.84 \\ 0.29 & -0.31 & -0.49 & 0.55 \end{bmatrix}$$

The best  $rank = 2$  approximation to  $\tilde{X}$  will be  $\tilde{X}^{*2} = \tilde{U}_2 \text{diag}(\tilde{d}_1, \tilde{d}_2) \tilde{V}_2'$ , where  $\tilde{U}_2$  is a matrix with the first two columns of  $\tilde{U}$ , and  $\tilde{V}_2$  is a matrix with the first two columns of  $\tilde{V}$ :

$$\tilde{X}^{*2} = \tilde{U}_2 \text{diag}(\tilde{d}_1, \tilde{d}_2) \tilde{V}_2' = \begin{bmatrix} -0.57 & 0.17 \\ 0.51 & 0.11 \\ -0.5 & -0.25 \\ 0.34 & -0.69 \\ 0.22 & 0.65 \end{bmatrix} \times \begin{bmatrix} 3.83 & 0 \\ 0 & 3.38 \end{bmatrix} \times \begin{bmatrix} 0.34 & -0.37 & -0.58 & 0.64 \\ -0.81 & 0.18 & 0.03 & 0.56 \end{bmatrix}$$

$$= \begin{bmatrix} -1.2 & 0.91 & 1.28 & -1.1 \\ 0.36 & -0.65 & -1.11 & 1.47 \\ 0.02 & 0.56 & 1.07 & -1.71 \\ 2.32 & -0.89 & -0.83 & -0.46 \\ -1.49 & 0.08 & -0.41 & 1.79 \end{bmatrix}$$

**Part f - NEED TO FINISH**

**Part d - standardized**

Note that

$$\tilde{\tilde{X}} = \begin{bmatrix} -0.63 & 0.96 & 1.38 & -0.49 \\ 0.63 & -0.24 & -0.15 & 1.34 \\ 0 & 0.96 & 0.61 & -1.1 \\ 1.26 & -1.43 & -0.92 & -0.49 \\ -1.26 & -0.24 & -0.92 & 0.73 \end{bmatrix}$$

The Singular Value Decomposition of  $\tilde{\tilde{X}}$  is:

$$\tilde{\tilde{U}} = \begin{bmatrix} -0.6 & 0 & 0.14 & -0.65 \\ 0.3 & 0.22 & 0.79 & 0.2 \\ -0.44 & -0.34 & -0.16 & 0.68 \\ 0.58 & -0.58 & -0.24 & -0.27 \\ 0.16 & 0.7 & -0.53 & 0.04 \end{bmatrix}, \quad \tilde{\tilde{D}} = \begin{bmatrix} 3.04 & 0 & 0 & 0 \\ 0 & 2.14 & 0 & 0 \\ 0 & 0 & 1.39 & 0 \\ 0 & 0 & 0 & 0.49 \end{bmatrix}$$

$$\tilde{\tilde{V}} = \begin{bmatrix} 0.36 & -0.69 & 0.56 & 0.29 \\ -0.63 & 0.13 & 0.19 & 0.74 \\ -0.6 & -0.17 & 0.49 & -0.61 \\ 0.33 & 0.69 & 0.64 & 0 \end{bmatrix}$$

Therefore, the principal component directions will be the columns of  $\tilde{\tilde{V}}$ , the principal components will be the inner products  $z_j = \langle x_i, v_j \rangle$ , the columns of:

$$z = \begin{bmatrix} -1.82 & -0.01 & 0.19 & -0.32 \\ 0.92 & 0.48 & 1.09 & 0.1 \\ -1.34 & -0.73 & -0.22 & 0.33 \\ 1.75 & -1.24 & -0.33 & -0.13 \\ 0.49 & 1.5 & -0.73 & 0.02 \end{bmatrix}$$

The “loadings” of the first principal component are the first column of the matrix  $\tilde{\tilde{V}}$ ,

$$\tilde{\tilde{V}}_1 = \begin{bmatrix} 0.36 \\ -0.63 \\ -0.6 \\ 0.33 \end{bmatrix}$$

**Part e - standardized**

The best  $rank = 1$  approximation to  $\tilde{\tilde{X}}$  will be  $\tilde{\tilde{X}}^{*1} = \tilde{\tilde{U}}_1 \text{diag}(\tilde{\tilde{d}}_1), \tilde{\tilde{V}}_1$ , where  $\tilde{\tilde{U}}_1$  is a matrix with first column of  $\tilde{\tilde{U}}$  and  $\tilde{\tilde{V}}_1$  is a matrix with the first column of  $\tilde{\tilde{V}}$ :

$$\begin{aligned}\tilde{X}^{*1} &= \tilde{U}_1 \text{diag}(\tilde{d}_1) \tilde{V}_1' = \begin{bmatrix} -0.6 \\ 0.3 \\ -0.44 \\ 0.58 \\ 0.16 \end{bmatrix} \times [3.04] \times \begin{bmatrix} 0.36 & -0.63 & -0.6 & 0.33 \end{bmatrix} \\ &= \begin{bmatrix} -0.65 & 1.15 & 1.09 & -0.61 \\ 0.33 & -0.58 & -0.55 & 0.31 \\ -0.48 & 0.85 & 0.8 & -0.45 \\ 0.63 & -1.11 & -1.05 & 0.58 \\ 0.18 & -0.31 & -0.29 & 0.16 \end{bmatrix}\end{aligned}$$

The best  $rank = 2$  approximation to  $\tilde{X}$  will be  $\tilde{X}^{*2} = \tilde{U}_2 \text{diag}(\tilde{d}_1, \tilde{d}_2) \tilde{V}_2'$ , where  $\tilde{U}_2$  is a matrix with the first two columns of  $\tilde{U}$ , and  $\tilde{V}_2$  is a matrix with the first two columns of  $\tilde{V}$ :

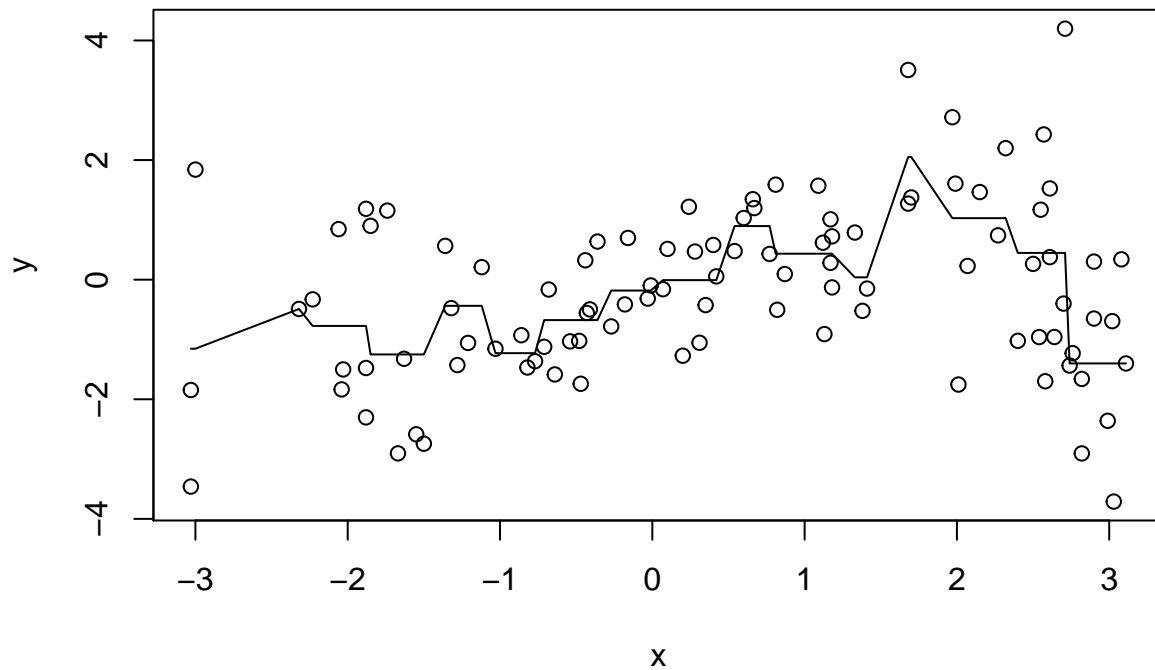
$$\begin{aligned}\tilde{X}^{*2} &= \tilde{U}_2 \text{diag}(\tilde{d}_1, \tilde{d}_2) \tilde{V}_2' = \begin{bmatrix} -0.6 & 0 \\ 0.3 & 0.22 \\ -0.44 & -0.34 \\ 0.58 & -0.58 \\ 0.16 & 0.7 \end{bmatrix} \times \begin{bmatrix} 3.04 & 0 \\ 0 & 2.14 \end{bmatrix} \times \begin{bmatrix} 0.36 & -0.63 & -0.6 & 0.33 \\ -0.69 & 0.13 & -0.17 & 0.69 \end{bmatrix} \\ &= \begin{bmatrix} -0.65 & 1.15 & 1.09 & -0.61 \\ 0 & -0.52 & -0.63 & 0.64 \\ 0.03 & 0.75 & 0.93 & -0.95 \\ 1.49 & -1.27 & -0.84 & -0.27 \\ -0.86 & -0.12 & -0.55 & 1.2 \end{bmatrix}\end{aligned}$$

**Part f - standardized - NEED TO FINISH**

## Problem 2

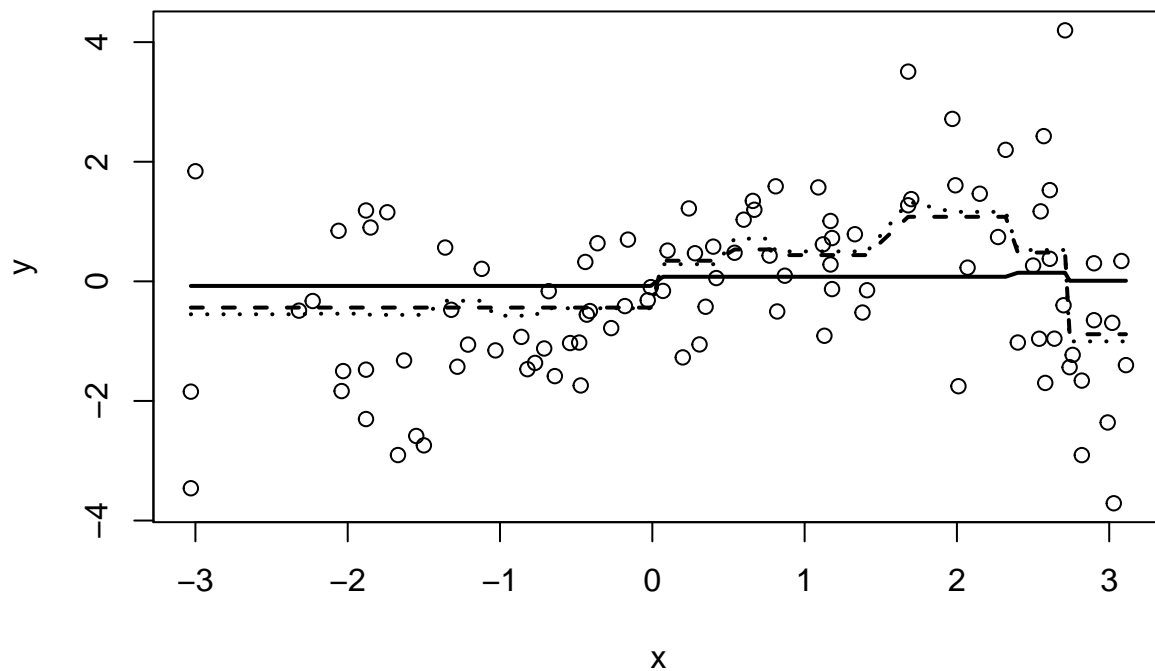
**Part a**

A scatterplot of the data and the fitted values from the OLS model are shown below.



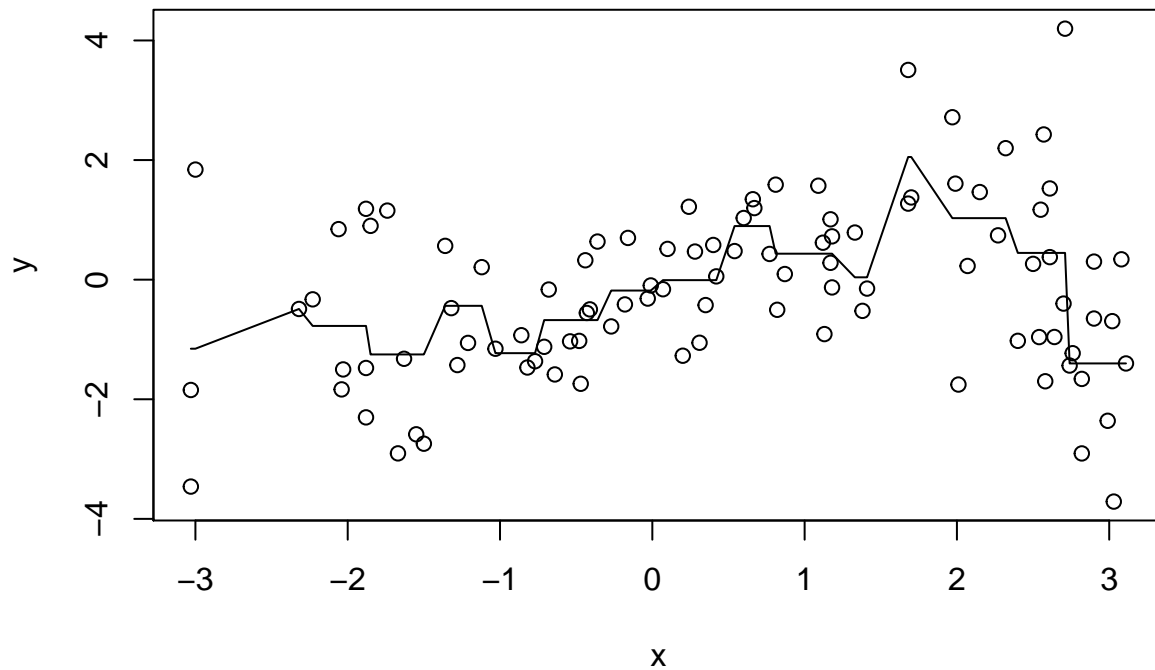
### Part b

The solid line in the plot below gives (approximately) the fitted values when there are 2 nonzero coefficients, the dashed line are fitted values when there are 4 nonzero coefficients, and the dotted line are fitted values for 8 nonzero coefficients.



### Problem 3

Note: This is Stat 602 HW3 from 2015 (problem 21)



#### Problem 4

Note: This is Stat 602 HW3 from 2015 (problem 22)

Note: There is some crossover with 502X HW2 Q15(a) here.

Part a

Part b

Part c

#### Problem 5

Note: This is Stat 602 HW3 from 2015 (problem 23)

Part a

Part b

#### Problem 6

Part a

The plots for this problem can be seen in Figure 1.



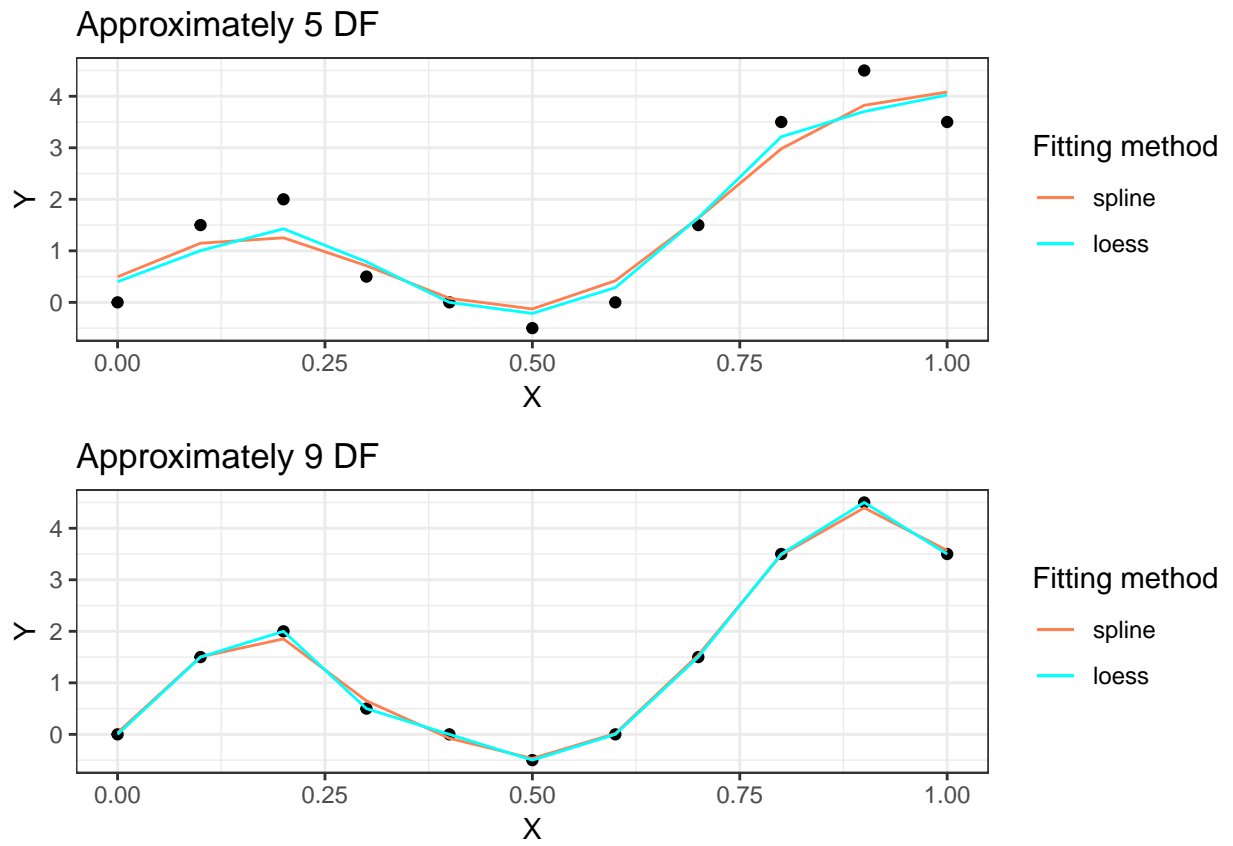


Figure 1: Plots for effective degrees of freedom 5 (top) and 9 (bottom) for Problem 6(a).

## Problem 7

kNN

elastic net

PCR

PLS

MARS (in earth)

## Problem 8

Part a

Part b

Part c

Part d